

A Bayesian hierarchical model for monthly maxima of instantaneous flow

Egil Ferkingstad

Science Institute, University of Iceland

and

Oli Pall Geirsson

Science Institute, University of Iceland

and

Birgir Hrafnkelsson

Faculty of Physical Sciences, University of Iceland

and

Olafur Birgir Davidsson

deCODE genetics

and

Sigurdur Magnus Gardarsson

Faculty of Environmental and Civil Engineering, University of Iceland

Monday 27th June, 2016

Abstract

We propose a comprehensive Bayesian hierarchical model for monthly maxima of instantaneous flow in river catchments. The Gumbel distribution is used as the probabilistic model for the observations, which are assumed to come from several catchments. Our suggested latent model is Gaussian and designed for monthly maxima, making better use of the data than the standard approach using annual maxima.

At the latent level, linear mixed models are used for both the location and scale parameters of the Gumbel distribution, accounting for seasonal dependence and covariates from the catchments. The specification of prior distributions makes use of penalised complexity (PC) priors, to ensure robust inference for the latent parameters. The main idea behind the PC priors is to shrink toward a base model, thus avoiding overfitting. PC priors also provide a convenient framework for prior elicitation based on simple notions of scale. Prior distributions for regression coefficients are also elicited based on hydrological and meteorological knowledge. Posterior inference was done using the MCMC split sampler, an efficient Gibbs blocking scheme tailored to latent Gaussian models. The proposed model was applied to observed data from eight river catchments in Iceland. A cross-validation study demonstrates good predictive performance.

Keywords: latent Gaussian models, extreme values, hydrology, penalised complexity priors

1 Introduction

A common and substantial problem in hydrology is that of estimating the return period of extreme floods. An accurate estimate of extreme floods is of interest in various circumstances, particularly with respect to important civil infrastructure. The design and construction of bridges and roads is often dependent on accurate understanding of river behavior during extreme events. Changes in land use, especially in the urban environment, create increasingly more impervious surfaces. This leads to larger and more frequent floods, putting more stresses on flood control structures, such as levees and dams. Climate change alters local precipitation patterns and magnitudes. This influences water resource management of reservoirs and rivers, affecting operation of hydroelectric power plants and river transport. The management, operation, and maintenance of this critical infrastructure relies on accurate flood predictions, including predictions for ungauged catchments based on data from gauged river catchments.

One of the first approaches to regional flood estimation was the *index flood method*, first proposed by Dalrymple [1960]. It was designed to deal with cases where little or no at-site data is available for flood assessment by borrowing strength from similar (e.g. neighboring) gauged catchments. The method consists of two main steps, namely, regionalization, which includes the identification of geographically and climatologically homogeneous regions, and the specification of a regional standardized flood frequency curve for a T -year return period. In Section 3 a mathematical formalization of the index flood method is used to motivate some of the elements of our proposed model.

The index flood method is still widely used today, and further developments of the method were presented in Hosking et al. [1985] and GREHYS [1996]. Starting with the work of Cunnane and Nash [1974], various Bayesian extensions have been proposed [Ros-

bjerg and Madsen, 1995, Kuczera, 1999, Martins and Stedinger, 2000]. Although these papers show the usefulness of Bayesian methods, they all derive rather directly from the classical index flood method, their main goal is usually to improve the estimation of the index flood coefficient, and they all rely solely on annual maxima. This work improves on the above studies in many important ways: The power relationship used to estimate the index flood coefficients is instead employed in the priors for the parameters of the Gumbel distribution, which we have chosen as the distribution for the observations. We use carefully chosen meteorological and topographical covariates, including catchment areas and covariates based on precipitation and temperature measurements, motivated by the work of Crochet [2012]. In summary, we believe that our work provides a coherent and comprehensive Bayesian model, making better use of the available data and prior knowledge.

We propose a Bayesian hierarchical model for monthly instantaneous extreme flow data from several river catchments. The topographical and climatic covariates facilitate the process of extrapolating the model to ungauged river catchments. Several novelties in statistical modeling and inference for flood data are presented here: We use monthly rather than yearly maxima, making better use of the available data. We use a latent Gaussian model (LGM, see e.g. Rue et al. [2009]) incorporating seasonal dependence, borrowing strength across months. The LGM allows the use of the computationally efficient MCMC split sampling algorithm [Geirsson et al., 2015], while still being sufficiently general to allow for realistic modeling. We use penalised-complexity priors [Simpson et al., 2014] for the hyperparameters of the model, which avoids overfitting, letting the prior knowledge together with the data decide the appropriate level of model complexity. We do a thorough prior elicitation for the regression coefficients of our model, making good use of available prior knowledge. To demonstrate that the proposed model predicts well for ungauged catchments, we perform a cross-validation study, where we leave river j out and predict based

on the model estimated from the other rivers except j , for each of the eight rivers.

We proceed as follows: Section 2 presents the data and the hydrological aspects of the problem. Section 3 introduces the full hierarchical model and provides explanations of the modelling assumptions, and a description of the posterior inference. Section 4 summarizes the results obtained from applying the model to the data. Finally, Section 5 contains the conclusions drawn from the study and some ideas for future research.

2 Data

2.1 Streamflow Data and River Catchments

The streamflow data consist of monthly maximum instantaneous discharges from eight river catchments in Iceland. Table 1 lists the identification number, name and the size of each catchment. Even though stations VHM45 and VHM204 have the same name (Vatnsdalsa), they correspond to different catchments. The time series were between 20 and 80 years long (in most cases between 40 and 60 years). Figure 1 shows the locations of the eight catchments.

Table 1: Characteristics of the catchments used in the study. The station identifications, river names and catchment areas were provided by the Icelandic Meteorological Office.

| Station | River | Area (km ²) | Station | River | Area (km ²) |
|---------|------------|-------------------------|---------|-------------|-------------------------|
| VHM10 | Svarta | 392 | VHM51 | Hjaltadalsa | 296 |
| VHM19 | Dynjandisa | 37 | VHM198 | Hvala | 195 |
| VHM26 | Sanda | 267 | VHM200 | Fnjoska | 1094 |
| VHM45 | Vatnsdalsa | 456 | VHM204 | Vatnsdalsa | 103 |

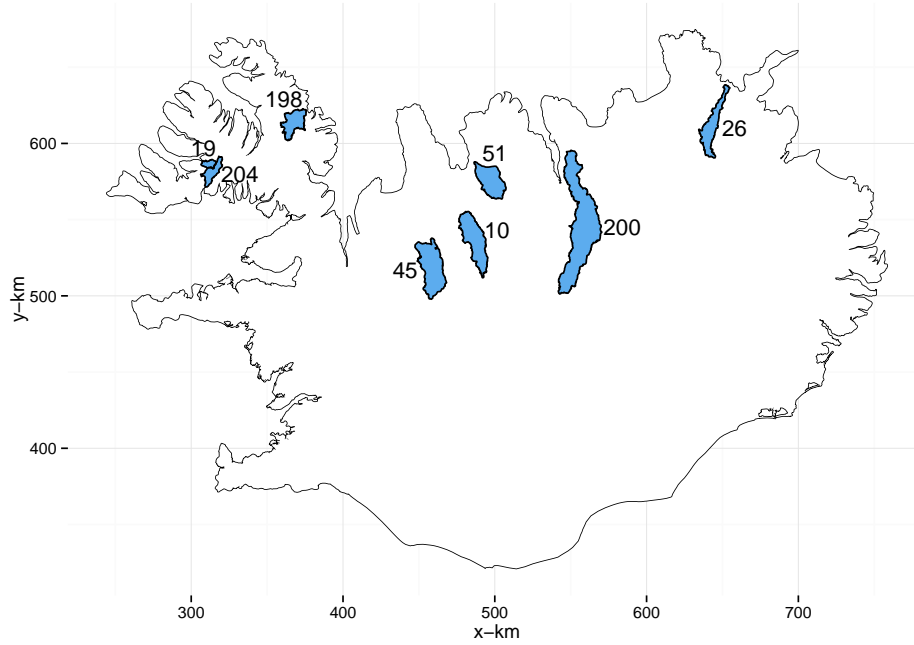


Figure 1: Locations of river catchments. Catchment boundaries are provided by the Icelandic Meteorological Office. Coastline is provided by Landsvirkjun, the National Power Company of Iceland.

Figure 2 shows the sample mean of the maximum monthly instantaneous flow for each river. The catchments have a seasonal behavior characterised by lower discharge during winter and higher discharge during spring/summer. The high discharge during spring/summer is mainly due to rising temperatures and snow melt, but the specific timing of the snow melt period varies somewhat for these catchments.

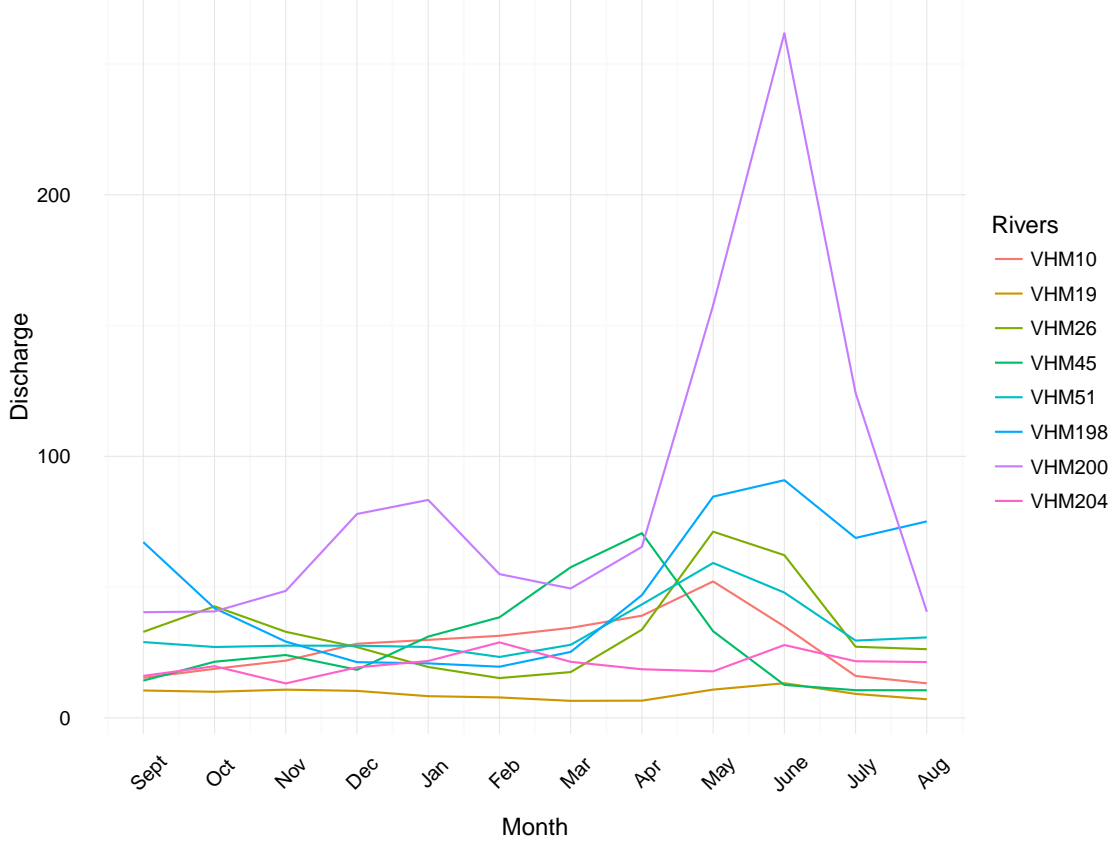


Figure 2: Sample means of maximum monthly flow in m^3/s for each river.

2.2 Topographical and Climatic Covariates

For each catchment, the following topographic and climatic covariates were considered for extrapolating to ungauged catchments:

Catchment area: The area of the river catchment in km^2 .

Average precipitation: The averaged monthly precipitation over the entire catchment.

To construct this covariate the precipitation on a 1 km by 1 km grid over the whole

of Iceland was obtained [Crochet et al., 2007], which was then integrated over the catchment area. Finally, the average over all years was found within each month.

Maximum daily precipitation: Daily precipitation over the catchment area within each month was acquired using the same method as for the average precipitation. The value corresponding to the day with the highest precipitation, cumulated over the catchment, was chosen, then the average over all years was found within each month.

Accumulated precipitation: The accumulated precipitation over the catchment since the start of the hydrological year (September). This covariate was potentially useful for explaining high discharge attributed to snow melt.

Average positive temperature: Temperature is available on the same grid as precipitation. These values were obtained in the same manner as the average precipitation within each month, with negative values truncated to zero.

Maximum positive temperature: These values were calculated in a similar way to the maximum precipitation values, with the difference being that negative temperature values were truncated to zero.

3 Models and Inference

3.1 Preliminary modeling and analysis

The Gumbel distribution is a common choice for extreme value data, due to its theoretical foundations [Coles et al., 2001]. We performed an Anderson–Darling goodness-of-fit test for the Gumbel distribution, for each river and month. The resulting p -values are shown

in Figure 3. The empirical distribution of the p -values is close to standard uniform, which suggests that the Gumbel distribution fits the observed data reasonably well.

We performed a preliminary analysis of the statistical relationship between maximum instantaneous flow and the topographical and meteorological factors described in Section 2. The preliminary analysis was carried out as follows. First, maximum likelihood (ML) estimates for both the location and scale parameters of the Gumbel distribution were obtained at all $J = 8$ rivers and every month m . We then fitted log-linear models where the ML estimates of the location and scale parameters, respectively, acted as the response, and all combinations of the aforementioned covariates are assessed. This preliminary analysis revealed a strongly significant log-linear relationship between the ML estimates of the location parameter and catchment area, average precipitation, maximum precipitation and accumulated precipitation. The analysis further showed a strong multicollinearity between average precipitation, maximum precipitation and accumulated precipitation. However, non-significant log-linear relationships were observed between the ML estimates and both average and maximum positive temperature. Based on these results and by using a step-wise log-linear model selection algorithm based on AIC score, it was decided to include both catchment area (x_1) and maximum daily precipitation (x_2) as predictive covariates for location parameters. Analogous results also hold for the scale parameter.

3.2 Description of the proposed hierarchical model

In this section, we present the proposed three-level Bayesian hierarchical model. At the data level, the observed maxima of instantaneous flow $y_{jm,t}$ for river j , month m , and year t is assumed to follow a Gumbel distribution:

$$y_{jm,t} \sim \text{Gumbel}(\mu_{jm}, \sigma_{jm}), \quad j = 1, \dots, J, \quad m = 1, \dots, M, \quad t = 1, \dots, T_{jm} \quad (1)$$

Let $\eta_{jm} = \log \mu_{jm}$ and $\tau_{jm} = \log \sigma_{jm}$. The linear model for η_{jm} is given by

$$\eta_{jm} = (\beta_0 + \beta_{0,m}^*)x_{0,jm} + (\beta_1 + \beta_{1,m}^*)x_{1,jm} + \cdots + (\beta_p + \beta_{p,m}^*)x_{p,jm} + \epsilon_{jm} \quad (2)$$

where the $x_{k,jm}$'s are centered log covariates (except $x_{0,jm} = 1$ for all j and m) and the random effect terms $\beta_{k,m}^*$ are given a prior enforcing seasonal behavior, described below. This model can be written in matrix form, as follows. Collect the covariates in the matrix \mathbf{X} , such that the first M rows of \mathbf{X} contain the covariates for river $j = 1$ over each of the M months, the next M rows contain the covariates for river $j = 2$, and so on. Let $(\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_p)$ denote the columns of \mathbf{X} , and let

$$\mathbf{Z}_k = \text{diag}(\mathbf{X}_k)(\mathbf{1}_J \otimes \mathbf{I}_M)$$

and

$$\mathbf{Z} = (\mathbf{Z}_0, \mathbf{Z}_1, \dots, \mathbf{Z}_p).$$

Further, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)'$, and let $\boldsymbol{\eta}$, $\boldsymbol{\beta}^*$ and $\boldsymbol{\epsilon}_\eta$ contain the η_{jm} , $\beta_{k,m}^*$ and ϵ_{jm} , ordered such that they line up with \mathbf{X} and \mathbf{Z} . Then we may write

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\beta}^* + \boldsymbol{\epsilon}_\eta.$$

The model for τ_{jm} is similar, with the same covariates, but different coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ and error term $\boldsymbol{\epsilon}_\tau$, and can be written in matrix form as

$$\boldsymbol{\tau} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\alpha}^* + \boldsymbol{\epsilon}_\tau.$$

To obtain a latent Gaussian model we must specify multivariate normal priors for the coefficients $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$. For $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ we fix $\boldsymbol{\mu}_\alpha$, $\boldsymbol{\mu}_\beta$, $\boldsymbol{\Sigma}_\alpha$ and $\boldsymbol{\Sigma}_\beta$ and set

$$\pi(\boldsymbol{\beta}) = \text{N}(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta), \quad \pi(\boldsymbol{\alpha}) = \text{N}(\boldsymbol{\alpha}|\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha),$$

where the choices of $\boldsymbol{\mu}_\alpha$, $\boldsymbol{\mu}_\beta$, $\boldsymbol{\Sigma}_\alpha$ and $\boldsymbol{\Sigma}_\beta$ are explained in Section 3.3. Let $\boldsymbol{\beta}_k^* = (\beta_{k,1}^*, \dots, \beta_{k,M}^*)$, $k = 0, \dots, p$ be the random intercepts ($k = 0$) or slopes ($k = 1, \dots, p$) of covariate k over the M months, and define $\boldsymbol{\alpha}_k^*$ similarly. We assume the following priors for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$, encoding seasonal dependence:

$$\pi(\boldsymbol{\beta}_k^*) = \text{N}(\boldsymbol{\beta}_k^* | \mathbf{0}, \psi_k^2 \mathbf{Q}^{-1}(\kappa)), \quad \pi(\boldsymbol{\alpha}_k^*) = \text{N}(\boldsymbol{\alpha}_k^* | \mathbf{0}, \phi_k^2 \mathbf{Q}^{-1}(\kappa)),$$

where ϕ_k^2 and ψ_k^2 are unknown variance parameters and $\mathbf{Q}(\kappa)$ is an $M \times M$ circular precision matrix that has the vector

$$s \cdot [1 \ f_1(\kappa) \ f_2(\kappa) \ f_1(\kappa) \ 1] = s \cdot [1 \quad -2(\kappa^2 + 2) \quad \kappa^4 + 4\kappa^2 + 6 \quad -2(\kappa^2 + 2) \quad 1]$$

on its diagonal band [Lindgren et al., 2011], where s is a constant ensuring that the inverse of the precision matrix is a correlation matrix. We have fixed κ to the value $\kappa = 1$, giving the prior correlation of 0.67 between neighboring months, which seems reasonable based on our prior knowledge. Note that s is a function of κ , e.g. for $\kappa = 1$, $s \approx 0.268$.

3.3 Priors for regression coefficients

We here present the priors for the regression coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. For each i , α_i and β_i will be given equal priors, since they enter the model in a similar way. The priors specified below are written in terms of β_i . As explained in Section 3.2, $\boldsymbol{\beta}$ should be given a multivariate normal prior. We will assume that the elements β_i are *a priori* independent, so we need to set independent normal priors for the individual coefficients β_i , $i = 0, \dots, p$.

We start by considering the coefficient β_1 corresponding to the logarithm of the size of the catchment area. First, note that negative values of β_1 make little sense, as this corresponds to a larger area giving lower maximum flows than a smaller area, other things being equal. To interpret the effects of varying positive values for β_1 , consider precipitation

events (rainy clouds) moving over the area. Each event will have a smaller spatial extent than the catchment area itself, when the catchment area is large; and a hypothetical increase of the catchment area corresponding to given precipitation event will lead to a smaller fraction of the area being covered by precipitation. This gives a “clustering effect”: smaller catchment areas will have a larger proportion covered by precipitation events than larger catchment areas. Since the value $\beta_1 = 1$ corresponds to a completely uniform distribution of precipitation (which is physically implausible), this means that β_1 is highly likely to be less than one. In other words, values $\beta_2 > 1$ correspond to an effect of area which increases larger than linearly, which is unrealistic for the abovementioned reasons.

Based on the above, we believe that the most sensible values for β_1 are in the interval $(0, 1)$. We propose that the normal prior density for β_1 is such that the probability of negative values is 0.05 and the probability of values greater than one is 0.05. These values result in a prior mean of 0.5 and a prior standard deviation of 0.304.

Considering the effect of precipitation given a fixed area, a similar line of argument can be given for the parameter β_2 corresponding to maximum daily precipitation: Higher maximum daily precipitation should result in higher flows, so the parameter should be positive. Also, $\beta_2 > 1$ is unrealistic for similar reasons as explained above for β_1 : natural clustering effects make super-linear effects of precipitation unlikely. Accordingly, β_2 is given the same $N(0.5, 0.304^2)$ prior as β_1 .

Since the data should provide good information for the intercept parameter β_0 , there is less of a need to specify an informative prior here. We have therefore chosen a normal density with mean zero and variance 10^4 as an uninformative prior for the intercept.

3.4 Penalised complexity priors for hyperparameters

In this section, we describe the selection of priors for the hyperparameters $\boldsymbol{\psi} = (\psi_0, \dots, \psi_p)$ and $\boldsymbol{\phi} = (\phi_0, \dots, \phi_p)$. We start by considering priors for $\boldsymbol{\psi}$. Note first that $\boldsymbol{\psi}$ can be regarded as a flexibility parameter: $\boldsymbol{\psi} = \mathbf{0}$ corresponds to a restricted model where we set $\boldsymbol{\beta}^* = \mathbf{0}$, i.e. the *base model*

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}_\eta$$

without correlated random effects. Simpson et al. [2014] provide a useful framework for selecting prior densities for flexibility parameters such as $\boldsymbol{\psi}$: penalised complexity (PC) priors. The ideas behind PC priors are thoroughly described in Simpson et al. [2014], but we give a short review here. PC priors are constructed based on four underlying principles. The first principle is Occam’s razor: we should prefer the simpler base model unless a more complex model is really needed. The second principle is using the Kullback-Leibler divergence (KLD) as a measure of complexity [Kullback and Leibler, 1951], where $\sqrt{2\text{KLD}}$ is used to measure the distance between the base model ($\boldsymbol{\psi} = \mathbf{0}$) and the more complex model corresponding to $\boldsymbol{\psi} > \mathbf{0}$ (the factor 2 is introduced for convenience, giving simpler mathematical derivations). The third principle is that of constant-rate penalisation, which is natural if there is no additional knowledge suggesting otherwise. This corresponds to an exponential prior on the distance scale $d = \sqrt{2\text{KLD}}$. Note that defining the prior on the distance scale implies that PC priors are invariant to reparameterization. The fourth and final principle is *user-defined scaling*, i.e. that the user should use (weak) prior knowledge about the size of the parameter to select the parameter of the exponential distribution. Simpson et al. [2014] provide both theoretical results and simulation studies showing the PC priors’ good robustness properties and strong frequentist performance.

We shall specify independent priors for each component ψ_k of $\boldsymbol{\psi}$ and each component

ϕ_k of ϕ . Note that this entails specifying separate base models for each component. While the ideal approach would be to specify an overall multivariate PC prior corresponding to the base model, we view this as beyond the scope of this article. It is easy to derive that the PC prior approach results in exponential priors for both the ψ_k and σ_η in this case, see Simpson et al. [2014] for details, so it only remains to specify the scaling, i.e. the choices of parameters of the respective exponential distributions.

The parameter ψ_0 is the standard deviation of the mean zero monthly intercepts $\beta_{0,m}^*$, representing the monthly deviations from the overall intercept β_0 . Since our model is on a logarithmic scale, the values $\beta_{0,m}^* = -4.61$ and $\beta_{0,m}^* = 4.61$ correspond to factors $\exp(-4.61) = 0.01$ and $\exp(4.61) = 100$, respectively, for $\exp(\beta_{0,m}^*)$. Accordingly, $(-4.61, 4.61)$ should be considered to be a wide 95% probability interval. The value of ψ_0 giving this interval is $\psi_0 = 2.35$. We take 2.35 as the 0.95 quantile of the prior for ψ_0 , giving a mean of 0.784 and a rate of 1.275 for the exponential prior for ψ_0 . A similar argument can be given for ϕ_0 , and we give it the same prior as ψ_0 .

Since ψ_1 , ψ_2 , ϕ_1 and ϕ_2 have similar roles in the model, they will given identical, independent, priors. We write in terms of ψ_1 below, with the understanding that the three other priors are identical. It is convenient to use a tail-area argument to specify the scaling. First, consider the sum of the “fixed effect” parameter β_1 and the “random effect” parameter $\beta_{1,m}^*$ for some month m . For the reasons described in Section 3.3, most of the prior mass of this sum should be between zero and one, but the addition of the random effects term will of course increase the variance, so the masses allowed below zero and above one should be larger than the 5% used in Section 3.3. We consider 10% prior mass below zero (and 10% above one) for $\beta_1 + \beta_{1,m}^*$ to give a relatively large mass outside the interval $(0, 1)$. This corresponds to a prior standard deviation of approximate 0.32 for each $\beta_{1,m}^*$. Since this is a high value, it should be in the upper tail of the prior for ψ_1 : We thus specify that 99%

of the mass of ψ_1 should be below the value 0.32, giving a rate of approximately 14.4 (and a mean of approximately 0.07) for the exponential prior for ψ_1 .

In lack of prior knowledge suggesting otherwise, we give equal priors to σ_η and σ_τ . The prior for σ_η can be specified in a more straightforward manner using a direct tail-area argument: Considering the scale of the problem, it seems highly likely that σ_η should be less than ten, so we put the 0.99-quantile of the exponential prior at the value ten. The result is a rate of 0.46 (and a mean of 2.17).

3.5 Posterior inference and computation

As latent models were imposed on both the location and scale parameters of the data density, approximation methods such as the integrated nested Laplace approximation [Rue et al., 2009] were inapplicable in our setting. Therefore, MCMC methods were necessary to make posterior inference. However, standard MCMC methods such as single site updating converged slowly and mixed poorly since many model parameters were heavily correlated in the posterior. For these reasons, all posterior inference was carried out by using the more efficient MCMC split sampler [Geirsson et al., 2015]. The MCMC split sampler is a two-block Gibbs sampling scheme designed for LGMs, where tailored Metropolis–Hastings strategies are implemented within in both blocks. The sampling scheme is well suited to infer LGMs with non-Gaussian data density where latent models are imposed on both the location and scale parameters.

The main idea of the MCMC split sampler is to split the latent Gaussian parameters into two vectors, called the “data-rich” block and the “data-poor” block. The data-rich block consists of the parameters that enter directly into the likelihood function, in our case the location parameters μ_{jm} and the scale parameters σ_{jm} , for $j = 1, \dots, J$ and $m = 1, \dots, M$. The data-poor block consists of the remaining parameters (in our case, including

the regression parameters and hyperparameters). An efficient block Gibbs sampling scheme can then be implemented by sampling from the full conditional distributions of each block. For the data-poor block, it turns out that the full conditional is multivariate Gaussian, so sampling can be done quickly using a version of the one-block sampler of Knorr-Held and Rue [2002]. The data-rich block can also be sampled efficiently, for details see Geirsson et al. [2015].

4 Results

The model described in Section 3 was fitted using the MCMC split sampler, with 30000 iterations, discarding a burn-in of 10000. Runtime on a modern desktop (Ivy Bridge Intel Core i7-3770K, 16GB RAM and solid state hard drive), was approximately one hour. All the calculations were done using R.

Figure 4 shows prior densities (in orange) together with posterior densities (light blue) for the regression coefficients β and α . The posteriors look close to being normally distributed. We see that the intercepts (Figures 4(a) and 4(d)) are well identified, with modes close to -5 and -4 , respectively, even though they have a vague prior. This is as expected, since the intercepts correspond to an overall, “average” level which should be relatively easy to infer. The posteriors for the regression coefficients β_1 and α_1 , corresponding to log catchment area, (Figures 4(b) and 4(e)), look similar, though the posterior for α_1 (in the model for log scale) is slightly wider. Both have a mode of around 0.75, and most of the posterior mass in the region between 0.5 and 1. Posteriors for β_2 and α_2 , corresponding to maximum daily precipitation (Figures 4(c) and 4(f)) are wider than those for β_1 and α_1 , with most of the mass in the region between 0.4 and 1.5. The posterior mode of β_2 is around 0.9, while the posterior mode of α_2 is close to 1.0.

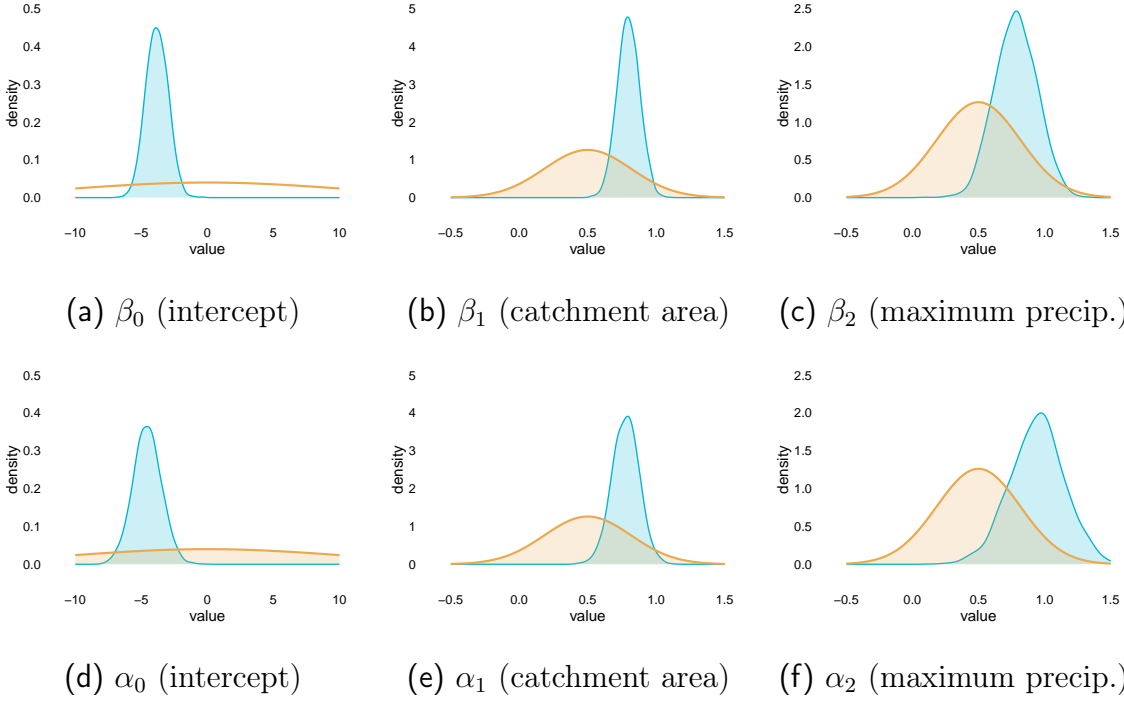


Figure 4: Prior (orange) and posterior densities (blue) for the regression coefficients.

Figure 5 shows prior and posterior densities for all eight hyperparameters of the model. We see that the hyperparameters for the random effects' standard deviations ψ_i and ϕ_i ($i = 0, 1, 2$) are all shrunk somewhat towards zero. However, the posterior mode is larger than zero for all hyperparameters, particularly for ϕ_1 , where there is very little mass close to zero. For the standard deviations ϕ_1, ϕ_2, ψ_1 and ψ_2 most of the posterior mass is between 0 and 0.1, while ϕ_0 and ψ_0 (corresponding to the random intercepts) have most of their posterior mass between 0 and 0.5. Posteriors for σ_η and σ_τ (the two residual noise standard deviations of the model) are well identified, even though they were given an very weakly informative prior. The posterior modes of σ_η and σ_τ are close to 0.5.

Figure 6 shows the seasonal effects, together with 80% pointwise credible intervals. It seems like there is some evidence for a seasonal effect for β_0^* (the intercept of the location model), and β_1^* and α_1^* (corresponding to catchment area), while this is not so clear for the other parameters. This is consistent with what was seen in Figure 5, particularly when comparing the posterior for ϕ_1 with the corresponding seasonal effect for α_1^* .

The left panels of Figure 7 show empirical cumulative distribution functions (CDFs) together with CDFs predicted from the model, for three randomly chosen river/month-combinations. The right panels show corresponding PP plots, i.e. the empirical CDF is plotted against the CDF predicted from the model for each river and each month. Uncertainty bands correspond to pointwise 95% credible intervals. The model seems to fit the data reasonably well.

Finally, we performed a cross-validation study, by leaving each river out in turn, estimating the full model based on the remaining seven rivers, and predicting for the left-out river. Figures 8 and 9 show the results for all eight rivers. Since the aim is to predict extremes, we do not consider prediction of the lower quantiles, but focus on the median and the 90th percentile. The limited number of data points (around 50) for each river-month

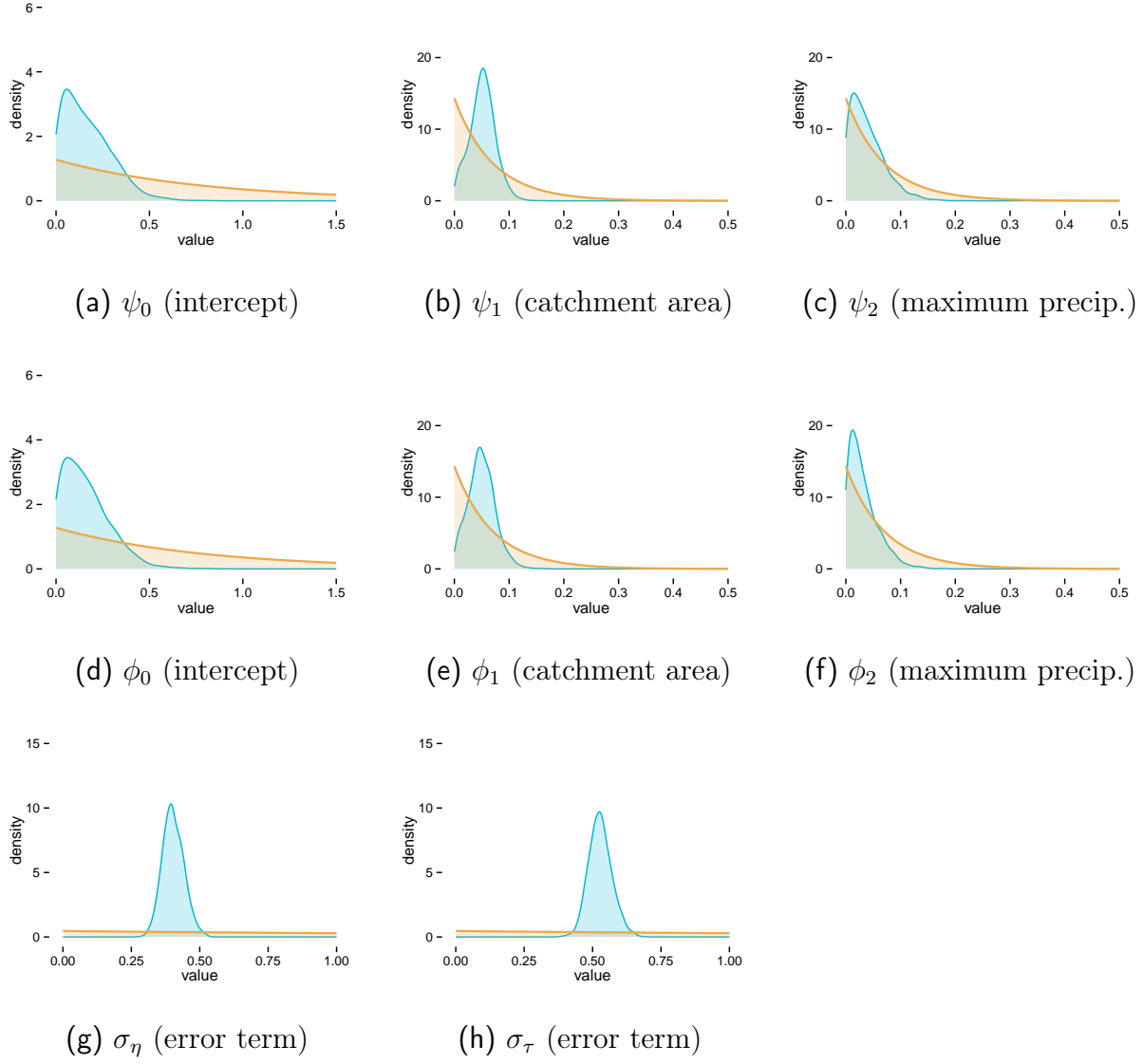
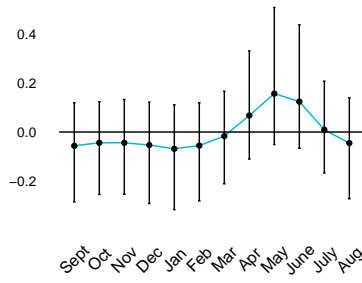
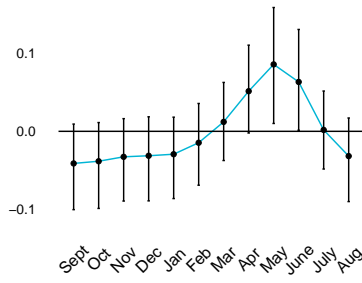


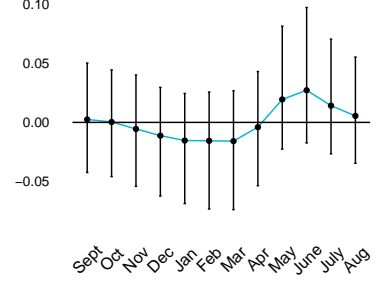
Figure 5: Prior (orange) and posterior (blue) densities of the hyperparameters.



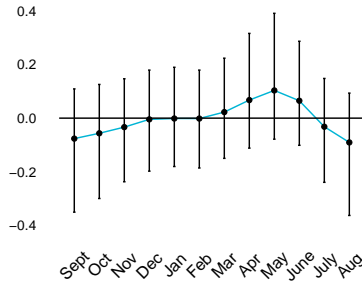
(a) β_0^* (intercept)



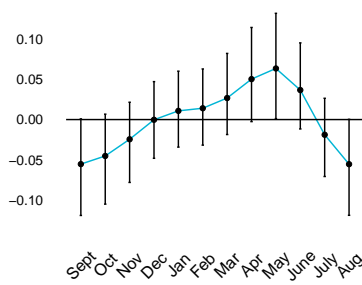
(b) β_1^* (catchment area)



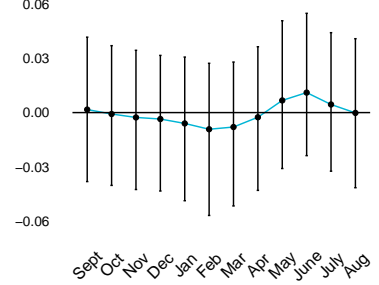
(c) β_2^* (maximum precip.)



(d) α_0^* (intercept)



(e) α_1^* (catchment area)



(f) α_2^* (maximum precip.)

Figure 6: Posterior mean and posterior 80% intervals for the seasonal effects.

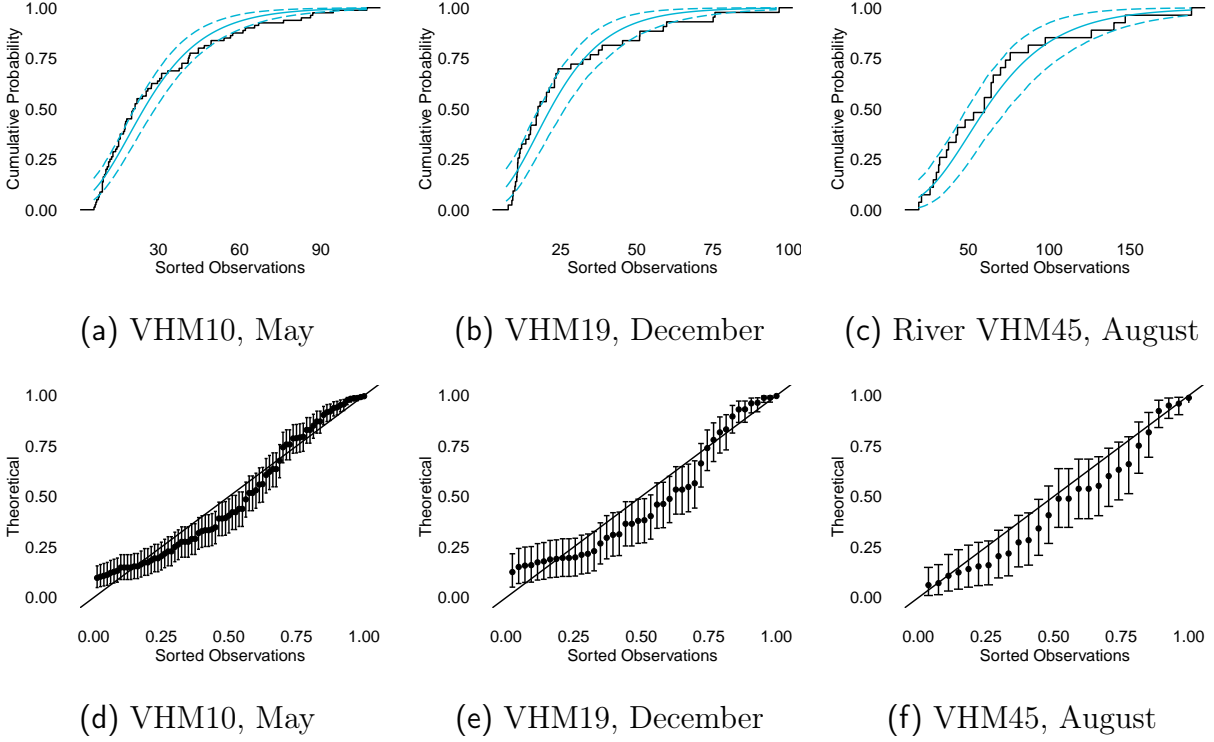


Figure 7: Model fit. The top panel shows predicted vs empirical cumulative distribution functions for three randomly chosen river-month combinations. The bottom panel shows probability-probability (pp) plots for the same river-month combinations, i.e. the empirical CDF is plotted against the CDF predicted from the model.

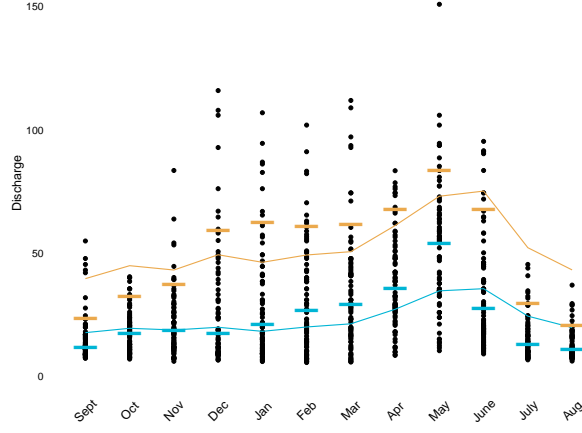
combination would make estimation of higher sample quantiles such as 0.95 or 0.99 too noisy.

The model seems to predict reasonably well overall, particularly when taking into account that the model was fitted based on only seven river catchments, and that these are a purely out-of-sample predictions based on sparse data. The worst prediction is for river VHM19, which is the smaller river catchment in our data set, and is also somewhat untypical, with smallest discharge levels overall. It is therefore perhaps not surprising that prediction fails somewhat here. For all the other rivers, however, the predictive accuracy is in our view about as good as can be expected.

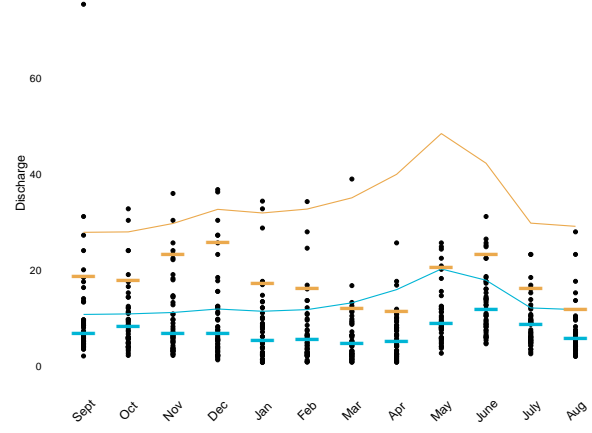
5 Conclusions

We have proposed a Bayesian hierarchical model for monthly maxima of instantaneous flow. Since the number of sites is often small (as in the data used here), the ability to borrow strength between months is very important. Rather than performing twelve (one for each month) independent linear regressions at the latent level, we fitted a linear mixed model using information jointly from all months and all sites. The use of penalised complexity priors was helpful, giving a good balance between prior information and sparse data. A thorough account of the prior elicitation for both regression coefficients and hyperparameters was given. We argue that the use of PC priors make hyperprior elicitation easier: the principle of user-defined scaling gives a useful framework for thinking about priors for hyperparameters in complex models.

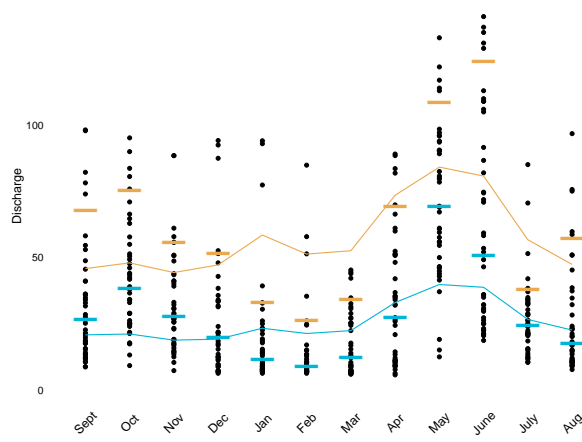
Based on a preliminary analysis, it was shown that the Gumbel distribution fits the data well in most cases. However, the generalised extreme value distribution is often selected as a model for block extrema, due to its theoretic basis and it containing the



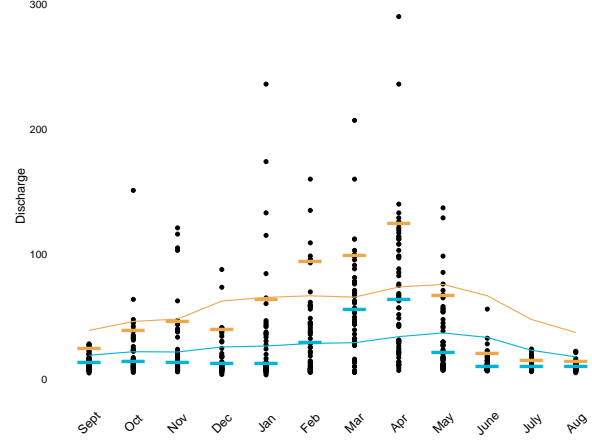
(a) VHM10



(b) VHM19

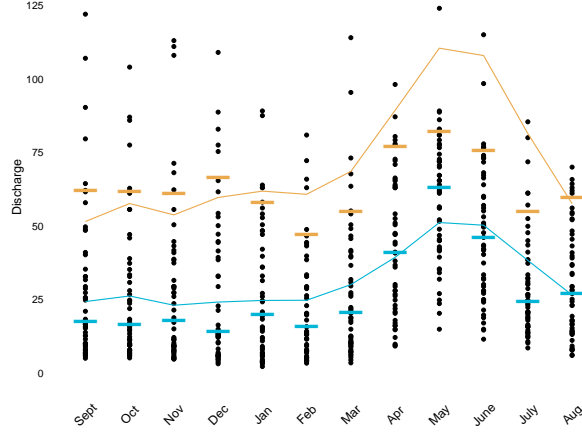


(c) VHM26

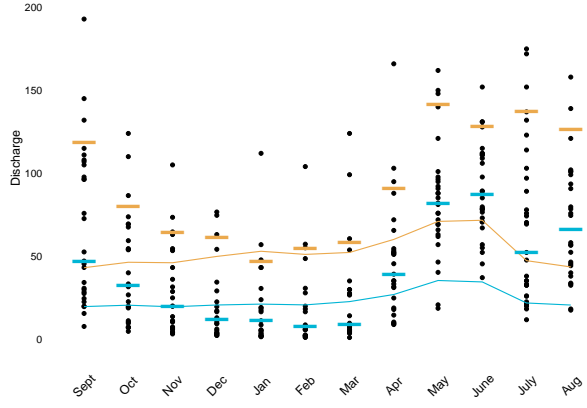


(d) VHM45

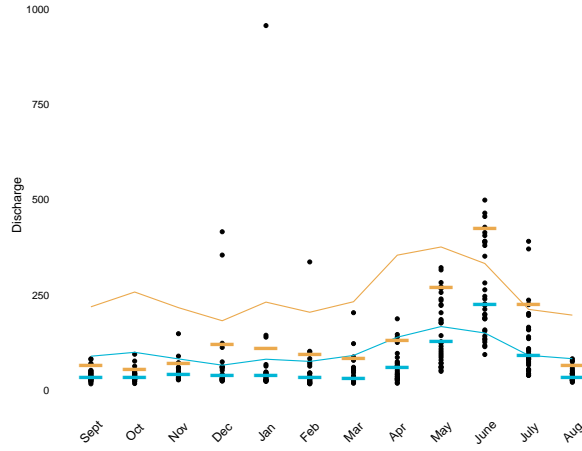
Figure 8: Predictive performance for rivers VHM10, VHM19, VHM26 and VHM45. The blue curves show predicted medians, while the orange curves show 90th percentile predictions. The blue bars show the data medians, while the orange bars show the 90th percentile of the data for each river. The black dots show the individual data points.



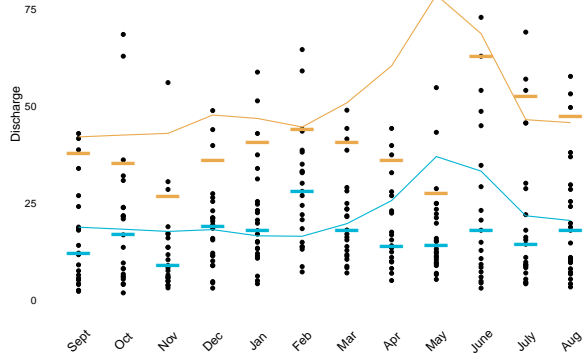
(a) VHM51



(b) VHM198



(c) VHM200



(d) VHM204Q

Figure 9: Predictive performance for rivers VHM51, VHM198, VHM200 and VHM204. The blue curves show predicted medians, while the orange curves show 90th percentile predictions. The blue bars show the data medians, while the orange bars show the 90th percentile of the data for each river. The black dots show the individual data points.

Gumbel distribution as a special case. Future research on models for monthly maxima of instantaneous flow should involve assuming the generalised extreme value distribution at the data level. Assuming the same shape parameter across months would be a sensible starting point. If that is not sufficient, then assuming that each month has its own shape parameter would be a sensible extension.

A crucial aspect of the proposed model is its capacity to predict monthly maxima of instantaneous flow at ungauged sites, provided that catchment covariates are available. The model could also be used to predict annual maxima of instantaneous flow at ungauged sites. The Bayesian approach allows for taking parameter uncertainty into account, while also helping to reduce uncertainty by using the regularising priors that are selected here. The result is reasonably good predictions compared to observed data.

Acknowledgements

We thank Håvard Rue, Andrea Riebler, Daniel Simpson and Philippe Crochet for many helpful comments and suggestions. The data was provided by the Icelandic Meteorological Office. The study was partly funded by the University of Iceland Research Fund.

References

- Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.
- Philippe Crochet. Estimating the flood frequency distribution for ungauged catchments using an index flood procedure. application to ten catchments in Northern Iceland. Technical report, Icelandic Meteorological Office, 2012.

- Philippe Crochet, Tómas Jóhannesson, Trausti Jónsson, Oddur Sigurðsson, Helgi Björnsson, Finnur Pálsson, and Idar Barstad. Estimating the spatial distribution of precipitation in Iceland using a linear model of orographic precipitation. *Journal of Hydrometeorology*, 8(6):1285–1306, 2007.
- C Cunnane and JE Nash. Bayesian estimation of frequency of hydrological events. *Mathematical models in hydrology*, 1, 1974.
- T Dalrymple. *Flood-frequency Analyses*. US Geological Survey, 1960.
- Óli Páll Geirsson, Birgir Hrafnkelsson, Daniel Simpson, and Helgi Sigurdarson. The MCMC split sampler: A block Gibbs sampling scheme for latent Gaussian models. *arXiv preprint arXiv:1506.06285*, 2015.
- Groupe de recherche en hydrologie statistique GREHYS. Presentation and review of some methods for regional flood frequency analysis. *Journal of Hydrology*, 186(1-4):63–84, 1996.
- JRM Hosking, James R Wallis, and Eric F Wood. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261, 1985.
- Leonhard Knorr-Held and Håvard Rue. On block updating in markov random field models for disease mapping. *Scandinavian Journal of Statistics*, 29(4):597–614, 2002.
- George Kuczera. Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resources Research*, 35(5):1551–1557, 1999.
- Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

- Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.
- Eduardo S Martins and Jerry R Stedinger. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(3):737–744, 2000.
- Dan Rosbjerg and Henrik Madsen. Uncertainty measures of regional flood frequency estimators. *Journal of Hydrology*, 167(1):209–224, 1995.
- Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical methodology)*, 71(2):319–392, 2009.
- Daniel P Simpson, Thiago G Martins, Andrea Riebler, Geir-Arne Fuglstad, Håvard Rue, and Sigrunn H Sørbye. Penalising model component complexity: A principled, practical approach to constructing priors. *arXiv preprint arXiv:1403.4630*, 2014.